Forward and Inverse Modelling of Lithospheric Deformation on Geological Timescales

B. J. P. Kaus, A. A. Popov, T. S. Baumann, A. E. Püsök, A. Bauville, N. Fernandez, M. Collignon

published in

NIC Symposium 2016
K. Binder, M. Müller, M. Kremer, A. Schnurpfeil (Editors)

Forschungszentrum Jülich GmbH,
John von Neumann Institute for Computing (NIC),
Schriften des Forschungszentrums Jülich, NIC Series, Vol. 48,
ISBN 978-3-95806-109-5, pp. 299.
http://hdl.handle.net/2128/9842

Forward and Inverse Modelling of Lithospheric Deformation on Geological Timescales

Boris J. P. Kaus[1,2,3], Anton A. Popov[1], Tobias S. Baumann[1], Adina E. Püsök[1], Arthur Bauville[1], Naiara Fernandez[1,4], and Marine Collignon[5]

[1] Institute of Geosciences, Johannes Gutenberg University Mainz,
Joh. Joachim Becherweg 21, 55128 Mainz, Germany
E-mail: {kaus, popov, baumann, puesoek, abauvill}@uni-mainz.de

[2] Center for Computational Sciences, JGU Mainz

[3] Center for Volcanoes and Atmosphere in Magmatic Open Systems (VAMOS), JGU Mainz

[4] now at: Bureau of Economic Geology, University of Texas, Austin, USA

[5] Centre for Earth Evolution and Dynamics, University of Oslo, Oslo, Norway

Geological processes such as mountain belt formation, subduction of tectonic plates and the development of sedimentary basins occur on a million-year timescale and involve rocks that have nonlinear visco-elasto-plastic material properties and experienced very large deformations. In order to simulate such processes in 3D, we developed a scalable parallel code, LaMEM, that employs a staggered finite difference discretisation combined with a marker and cell approach. Here, we describe the numerical approach and discuss some case studies in which we employed the code (i) to study the physics of crustal scale folding and faulting, (ii) to understand how continental collision might result in mountain belt and plateau formation, (iii) how it can be combined with an inversion strategy to constrain the rheology of the crust and lithosphere.

1 Introduction

Computational geodynamics uses numerical modelling to understand fundamental geoscientific questions such as: Why do we have plate tectonics on Earth and not on other planets? How did mountain belts such as the Alps and the Himalaya form and what is the role of erosion in this? At the same time, computational geodynamics also addresses more practical questions related to the evolution of so-called fold-and-thrust-belts and salt structures, which are closely linked with the majority of the world's oil reservoirs. Numerically, the problems are challenging as the rheology of rocks varies from elasto-plastic at low temperatures (close to the Earth surface), to viscous at higher temperatures (deeper in the lithosphere and mantle). In addition, the codes should be able to handle very large strains (even after plastic material failure has occurred). As geological processes are slow, inertial terms are negligible and one has to solve the (incompressible) Stokes equations but with visco-plastic or visco-elasto-plastic rheologies, which results in an elliptic system of equations with strongly varying coefficients (that can be 6 orders of magnitude or more over several grid cells). Over the last few years, we have developed a new code (LaMEM - Lithosphere and Mantle Evolution Model), which fulfils these requirements and has a range of multigrid preconditioners combined with Newton iterations for nonlinearities to solve the resulting equations. In addition, we have coupled the code with a Monte-Carlo inversion approach, in order to better constrain the mechanical structure of active mountain belts.

We will describe the mathematical background and numerical implementation of the forward and inverse modelling approach in Sec. 2, and discuss some modelling results in Sec. 3.

2 Numerical Approach

2.1 Mathematical Approach

We solve the coupled system of momentum, mass, and energy conservation equations, respectively, with velocity (v_i), pressure (p) and temperature (T) as primary unknowns:

$$\frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial p}{\partial x_i} + \rho g_i = 0 \tag{1}$$

$$\frac{1}{K}\frac{Dp}{Dt} - \alpha \frac{DT}{Dt} + \frac{\partial v_i}{\partial x_i} = 0 \tag{2}$$

$$\rho C_p \frac{DT}{Dt} = \frac{\partial}{\partial x_i}\left(\lambda \frac{\partial T}{\partial x_i}\right) + H. \tag{3}$$

Here $x_i (i = 1, 2, 3)$ denotes Cartesian coordinates, $\tau_{ij} = \sigma_{ij} + p\delta_{ij}$ is the Cauchy stress deviator, ρ density, g_i the gravity acceleration vector, K the bulk modulus, α the thermal expansion coefficient, C_p the specific heat, λ thermal conductivity, H volumetric heat source, and D/Dt stands for the material time derivative, respectively. The visco-elasto-plastic constitutive equation for the deviatoric stress is given by:

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}^{el}_{ij} + \dot{\varepsilon}^{vs}_{ij} + \dot{\varepsilon}^{pl}_{ij} = \frac{\overset{\diamond}{\tau}_{ij}}{2G} + \dot{\varepsilon}^{vs}_{II}\frac{\tau_{ij}}{\tau_{II}} + \dot{\varepsilon}^{pl}_{II}\frac{\tau_{ij}}{\tau_{II}}, \tag{4}$$

where $\dot{\varepsilon}_{ij} = \frac{1}{2}(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}) - \frac{1}{3}\frac{\partial v_k}{\partial x_k}\delta_{ij}$ is the deviatoric strain rate tensor, $\dot{\varepsilon}^{el}_{ij}$, $\dot{\varepsilon}^{vs}_{ij}$, $\dot{\varepsilon}^{pl}_{ij}$ are the elastic, viscous and plastic components, respectively, $\overset{\diamond}{\tau}_{ij} = \frac{\partial \tau_{ij}}{\partial t} + \tau_{ik}\omega_{kj} - \omega_{ik}\tau_{kj}$ is the Jaumann objective stress rate, $\omega_{ij} = \frac{1}{2}(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i})$ is the spin tensor, G is the elastic shear modulus, and the subscript $_{II}$ denotes the square root of the second invariant of a corresponding tensor, e.g. for the deviatoric stress $\tau_{II} = \left(\frac{1}{2}\tau_{ij}\tau_{ij}\right)^{1/2}$. The magnitude of the viscous creep strain rate is subdivided into diffusion ($\dot{\varepsilon}_l$), and dislocation ($\dot{\varepsilon}_n$) components:

$$\dot{\varepsilon}^{vs}_{II} = \dot{\varepsilon}_l + \dot{\varepsilon}_n = A_l\, \tau_{II} + A_n\, (\tau_{II})^n\,, \tag{5}$$

where n is the stress exponent of the dislocation creep, and the pre-exponential factor (A) of each creep mechanism is defined by:

$$A_l = B_l\, \exp\left[-\frac{E_l + pV_l}{RT}\right], \quad A_n = B_n\, \exp\left[-\frac{E_n + pV_n}{RT}\right]. \tag{6}$$

Here B, E, and V denote the creep constant, activation energy, and activation volume, respectively, of the corresponding creep mechanism, and R is the gas constant. The magnitude of the plastic strain rate ($\dot{\varepsilon}^{pl}_{II}$) is determined by enforcing the Drucker-Prager yield criterion:

$$\tau_{II} \leq \tau_Y = \sin(\phi)\, p + \cos(\phi)\, c, \tag{7}$$

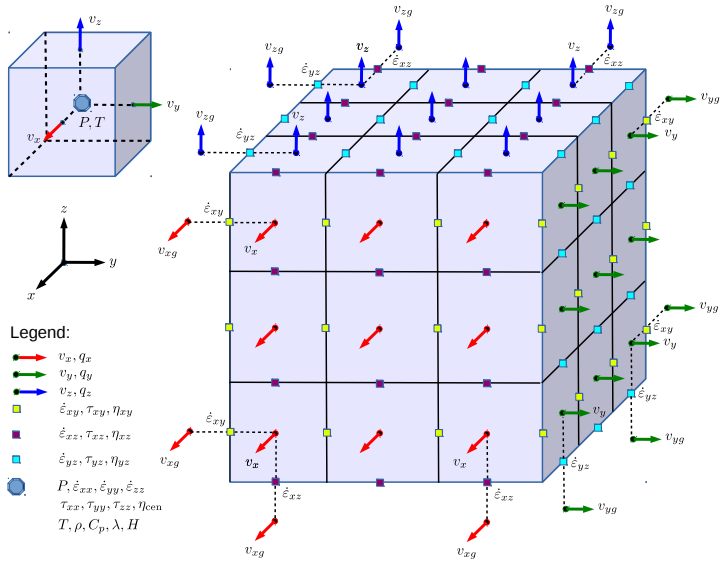

Figure 1. Staggered grid finite difference spatial discretisation, illustrating how the variables are ordered.

where ϕ is the friction angle, and c is the cohesion. The volumetric heat source includes shear heating, controlled by efficiency parameter $0 \leq \chi \leq 1$, and the radiogenic heat (A):

$$H = \chi \tau_{ij} \left(\dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^{el} \right) + \rho A. \tag{8}$$

2.2 Numerical Formulation

We discretise the conservation Eqs. 1–3 in space using staggered grid finite differences[1] as it is a low-order but stable discretisation for (nearly) incompressible fluid flow (see Fig. 1 for grid layout). To achieve scalability on massively parallel machines we use the distributed arrays (DMDA) and iterative solvers (KSP, SNES) from the PETSc library[2].The free surface is implemented using a so-called sticky air approach, which assigns a relatively low but nonzero viscosity to the air phase, together with an appropriate stabilisation method to allow for sufficiently large time steps[3,4]. The topography of the free surface is explicitly tracked by an internal 2D grid that covers the entire domain.

We employ a Marker And Cell (MAC) method[1] to track material properties and implement material advection in an Eulerian kinematical framework. To prevent spurious clustering of the material particles (markers) we use a combination of a 4^{th}-order Runge-Kutta method with a conservative velocity interpolation scheme[5]. During the advection, the elastic history stresses from previous time step (τ_{ij}^n) are corrected on the markers to account for the rigid-body rotation, and then interpolated on the edge and cell control volumes (Fig. 1) using the distance-based averaging[4] to obtain the effective strain rates:

$$\dot{\varepsilon}_{ij}^* = \dot{\varepsilon}_{ij} + \frac{\tau_{ij}^*}{2G\Delta t}, \quad \tau_{ij}^* = \tau_{ij}^n + \Delta t \left(w_{ik} \tau_{kj}^n - \tau_{ik}^n w_{kj} \right). \tag{9}$$

The second invariant of the effective strain rate is computed by cross-interpolation and averaging of the missing data (squares of the corresponding components) between all the control volumes.

The effective viscosity (η^*) and the updated deviatoric stresses (τ_{ij}) are computed from the effective strain rates, using the standard quasi-viscous expression:

$$\tau_{ij} = 2\eta^* \dot{\varepsilon}^*_{ij}, \quad \eta^* = \min\left[\left(\frac{1}{G\Delta t} + \frac{1}{\eta_l} + \frac{1}{\eta_n}\right)^{-1}, \frac{\tau_Y}{2\dot{\varepsilon}^*_{II}}\right]. \tag{10}$$

Here, the individual creep viscosities are defined by

$$\eta_l = \tfrac{1}{2}\left(A_l\right)^{-1}, \quad \eta_n = \tfrac{1}{2}\left(A_n\right)^{-\frac{1}{n}} \left(\dot{\varepsilon}^*_{II}\right)^{\frac{1}{n}-1}. \tag{11}$$

The discretised coupled system of nonlinear algebraic equations is solved at each time step using the preconditioned Jacobian-Free Newton-Krylov (JFNK) method with line-search as implemented in the PETSc SNES nonlinear solver framework[2]:

$$\mathbf{A}^{-1}\mathbf{J}\left(\mathbf{x}_k\right)\delta\mathbf{x}_k = -\mathbf{A}^{-1}\mathbf{r}\left(\mathbf{x}_k\right), \quad \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha\,\delta\mathbf{x}_k, \tag{12}$$

where $\mathbf{r}$ and $\mathbf{x}$ are the coupled residual and solution vectors, respectively, $\delta\mathbf{x}$ is the iterative correction vector, k is the iteration index, and α the line-search step length. The Jacobian ($\mathbf{J}$) is defined implicitly by a matrix-vector product approximated by finite differencing. The preconditioning matrix ($\mathbf{A}$) is obtained by discretising the conservation Eqs. 1–3 using the current effective viscosity and by ignoring the coupling terms between the Stokes block and the energy equation. $\mathbf{J}$ and $\mathbf{A}$ are given by:

$$\mathbf{J}\mathbf{y} \approx \frac{\mathbf{r}\left(\mathbf{x} + h\,\mathbf{y}\right) - \mathbf{r}\left(\mathbf{x}\right)}{h}, \quad \mathbf{A} = \begin{pmatrix} \mathbf{K} & \mathbf{G} & 0 \\ \mathbf{D} & \mathbf{C} & 0 \\ 0 & 0 & \mathbf{E} \end{pmatrix}, \tag{13}$$

where h is the perturbation parameter, $\mathbf{y}$ is an arbitrary vector to be multiplied with the Jacobian, $\mathbf{K}$, $\mathbf{C}$, and $\mathbf{E}$ denote the stiffness matrices of the velocity, pressure, and temperature blocks, respectively, $\mathbf{G}$ is the pressure gradient matrix and $\mathbf{D}$ is the velocity divergence matrix.

To achieve optimal scalability of the linear solver we employ a multigrid method to approximately invert the Stokes block in the preconditioning matrix. The coarse grid operators for the k-th level are obtained algebraically via Galerkin coarsening: process $\mathbf{A}_k = \mathbf{R}^k_{k+1}\ \mathbf{A}_{k+1}\ \mathbf{P}^{k+1}_k$. We have incorporated custom restriction ($\mathbf{R}$) and prolongation ($\mathbf{P}$) operators suitable for the staggered grid discretisation[6] into the PETSc multigrid framework. The multigrid preconditioner is implemented in either a coupled form, using simultaneous coarsening of the velocity and pressure blocks, or in a block triangular form, in which coarsening is applied only to the velocity matrix. Accordingly, the resulting preconditioners are referred to as coupled ($\mathbf{A_c}$) or uncoupled ($\mathbf{A_u}$)

$$\mathbf{A_c} = \begin{pmatrix} \mathbf{K} & \mathbf{G} \\ \mathbf{D} & -\frac{1}{\eta^*}\mathbf{I} \end{pmatrix}, \quad \mathbf{A_u} = \begin{pmatrix} \mathbf{K} & \mathbf{G} \\ 0 & -\frac{1}{\eta^*}\mathbf{I} \end{pmatrix}. \tag{14}$$

In both cases, we approximate pressure Schur complement by the inverse viscosity matrix.

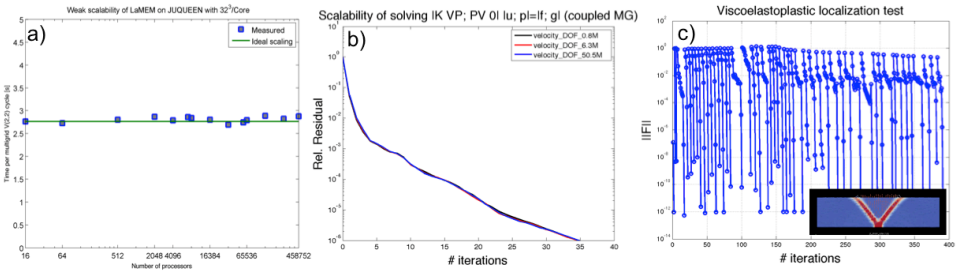

Figure 2. a) Weak scalability test of LaMEM for a single V-cycle with 32^3 gridpoints/processor. The largest simulation has $>$ 15 billion grid points. b) Convergence of a solver step for a variable viscosity falling block test with viscosity contrast 10^3, demonstrating that our coupled multigrid preconditioner results in resolution independent convergence (KSP: GMRES, with GMG and Chebyshev/Jacobi(3,3) smoothers & GMRES/GAMG as coarse grid solver). c) convergence of the nonlinear residual for various time steps for the viscoelastoplastic shear localisation setup shown in the inset.

2.3 Scalability Tests

We performed a number of tests to verify the scalability of LaMEM on JUQUEEN. A weak scalability test shows nearly perfect scalability for a single multigrid step up to the whole machine size (Fig. 2a). Our coupled multigrid preconditioner gives convergence behaviour that is resolution independent for a variable viscosity setup with a viscosity contrast of 1000 and 10 falling spheres, which was demonstrated to be a realistic test setup for geodynamic problems (Fig. 2b)[7]. A typical example of convergence during a visco-elasto-plastic shear localisation test during several time steps shows that the Newton solver results in rapid convergence once the initial residual has been reduced sufficiently (Fig. 2c).

3 Application Examples

3.1 Crustal-Scale Folding and Faulting

Under compression, crustal rocks can either fault or deform through folding. Folding results in quasi-regular structures, but it was incompletely understood how such structures grow, particularly in 3D. We therefore performed a systematic study and could demonstrate, using a combination of scaling laws and 3D simulations, that the wavelength of folds is mainly controlled by the effective viscosity structure of the crust that is deformed above a weaker salt layer. Lateral growth of folds results in linking of individual segments in a way that is quite similar to those observed in the Zagros (Iran)[8]. Yet, in some locations in the Zagros, salt crops out at the surface. Geological arguments suggest that these salt outcrops are salt diapirs sourced at a deep salt level, but that they were shallowly buried or exposed at the surface before the onset of collision. The formation of such salt structures is controlled by the speed with which sediments are deposited on top of the salt[9]. As such pre-existing salt structures form large-scale heterogeneities, they might affect the folding process. In order to understand that, we performed simulations in which we added the observed salt structure spacing in the Zagros to the initial model setup. Results show that this indeed localises deformation and result in folds with larger amplitudes, even though

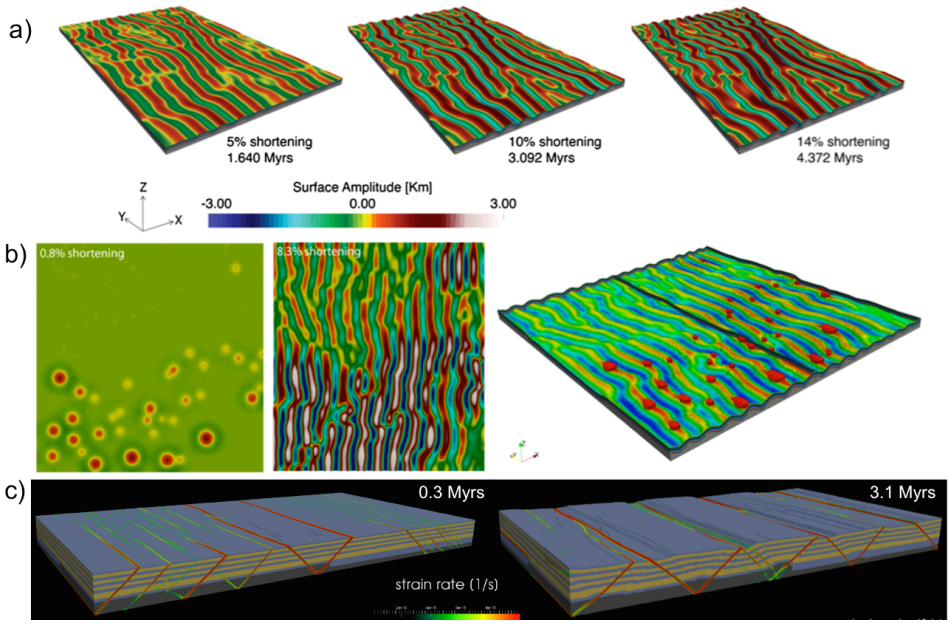

Figure 3. Examples of fold-and-thrust belt simulations: a) 3D fold pattern formation that develops once a system of multilayered crustal rocks above a weaker salt layer is compressed. Random heterogeneities are added to initiate the instability. Yet, fold spacing is mainly controlled by the material properties of the rocks. b) If pre-existing salt diapirs are present before the onset of collision, they affect the fold amplitudes but not their spacing. c) example of brittle fault formation and interaction during compression.

the fold spacing is unaffected (Fig. 3b)[10]. In additional work, we studied the effect of erosion on folding patterns and could demonstrate that erosion by itself does not affect fold spacing[11] dramatically, even though it might result in a faster growth of the structures and affect the manner in which folds link laterally[12].

Yet, folding is not the only response to compression. If crustal rocks are relatively cold, or have no or only few mechanically weak layers, fault zones will develop. In order to understand how this process works, we performed simulations of a brittle crust, which show that there is significant lateral interaction between fault zones, which ultimately develop larger scale structures (Fig. 3c) .

3.2 Continental Collision and Plateau Formation

The collision of India with Asia resulted in the largest mountain belt on Earth, the Himalaya, but also in the Tibetan plateau, which has an average height ($\sim$ 5km) that is more than the highest mountain in Europe. A similar plateau exists in the Andes (Altiplano), whereas the Alps does not have one. Why does a plateau form in some cases and not in others? In order to understand this, we performed systematic simulations using a setup in which an oceanic plate subducts underneath an overriding continental plate, followed by continent-continent collision. The results show that in order to form a plateau, we need to

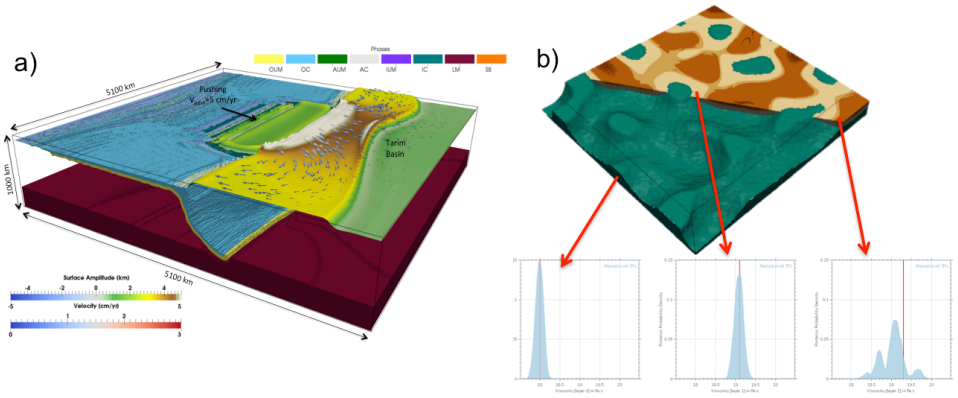

Figure 4. a) Simulation to study how subduction followed by continental collision results in the formation of a mountain belt and a high-elevation plateau (for example the Tibetan plateau)[13]. b) Example of geodynamic inverse modelling of salt-tectonics. During the inversion process, we assume that the geometry is known and we vary the density and viscosity of each of the layers. The modelled surface velocity and gravity anomalies are compared with the "observed" ones (in this case taken from a synthetic setup with known parameters). Results give a probability density function for each of the inversion parameters.

have sufficiently large convergence velocities, strong parts in the overriding lithosphere and a sufficiently large viscosity of the lithosphere (Fig. 4). We could demonstrate that there are 4 different types of plateaus that can form, which are controlled by two non-dimensional numbers, and which can be used to explain the observations[13].

3.3 Constraining the Rheology of the Lithosphere through Geodynamic Inverse Modelling

The largest uncertainty in performing geodynamic models comes from our imprecise knowledge of the effective viscosity of rocks, which is typically measured in the laboratory by deforming small rocks samples. Rather than relying on these data, we developed a new approach that couples lithospheric-scale geodynamic models with geophysical observations such as the GPS velocities with which plates deform, the topography of the lithosphere and measured gravity anomalies. During the inversion process, we automatically adopt the input parameters until a low misfit is obtained. As it is unclear whether a single global minimum exists, we employ a Monte Carlo based method in combination with geometrical constraints, which is able to deal with multiple local minima[14]. We could demonstrate with an analytical solution that this geodynamic inversion gives unique results for a rising sphere example, which is a significant improvement over earlier gravity-only inversions that has non-unique results. In a next step, we performed synthetic tests using linear viscous 3D models and showed that the method is able to retrieve the material parameters of layers that contribute to the large-scale dynamics of the model[14]. More recently, we could demonstrate that it also works for fully non-linear and temperature dependent rheologies that include plastic yielding[15], which implies that it can be employed to constrain the rheology of the lithosphere. A first application to the India-Asia collision zone shows that the viscosity of the Indian lithosphere must be rather large, but also that the

viscosity beneath Asia can be less well constrained given the current data[15]. Overall, our results suggest that the geodynamics inversion is a very promising new research direction that will give new insights in our understanding of the physics of the lithosphere.

Acknowledgements

This work was sponsored by the European Research Council under Starting Grant 258830. The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN[16] at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF).

References

1. F. Harlow, J. Welsh, *Numerical calculation of time-dependant viscous incompressible flow of fluid with free surface*, Phys. Fluids **8**, 2182–2189, 1965.
2. S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, K. Rupp, B. F. Smith, S. Zampini, H. Zhang, *PETSc users manual*, Technical Report ANL-95/11–Revision 3.6, Argonne National Laboratory, 2015.
3. B. J. P. Kaus, H. Mühlhaus, D. A. May. *A stabilization algorithm for geodynamic numerical simulations with a free surface.* Phys Earth Planet In **181**, 12–20, 2010.
4. T. Duretz, D. A. May, T. V. Gerya, P. J. Tackley, *Discretization errors and free surface stabilization in the finite difference and marker–in–cell method for applied geodynamics: A numerical study*, Geochem. Geophys. Geosyst. **12**, Q07004, doi:10.1029/2011GC003567, 2011.
5. H. Wang, R. Agrusta, J. van Hunen, *Advantages of a conservative velocity interpolation (CVI) scheme for particle-in-cell methods with application in geodynamic modeling*, Geochem. Geophys. Geosyst. **16**, 2015–2023, 2015.
6. M. Cai, A. Nonaka, J. B. Bell, B. E. Griffith, A. Donev, *Efficient Variable-Coefficient Finite-Volume Stokes Solvers*, Commun. Comput. Phys. **16**, 1263–1297, 2014.
7. D. A. May, J. Brown, L. Le Pourhiet, *pTatin3D: High-Performance Methods for Long-Term Lithospheric Dynamics*, SC14 , 274–284, 2014.
8. N. Fernandez, B. J. P. Kaus, *Fold interaction and wavelength selection in 3D models of multilayer detachment folding*, Tectonophysics **632**, 199–217, 2014.
9. N. Fernandez, B. J. P. Kaus, *Pattern formation in 3-D numerical models of down-built diapirs initiated by a Rayleigh-Taylor instability*, Geophys. J. Int. **202**, 1253–1270, 2015.
10. N. Fernandez, B. J. P. Kaus, *Influence of pre-existing salt diapirs on 3D folding patterns*, Tectonophysics **637**, 354–369, 2014.

11. M. Collignon, B. J. P. Kaus, D. A. May, N. Fernandez, *Influences of surface processes on fold growth during 3-D detachment folding*, Geochem. Geophys. Geosyst. **15**, 3281–3303, 2014.
12. M. Collignon, N. Fernandez, B. J. P. Kaus, *Influence of surface processes and initial topography on lateral fold growth and fold linkage mode*, Tectonics **34**, 1622–1645, 2015.
13. A. E. Püsök, B. J. P. Kaus, *Development of topography in 3-D continental-collision models*, Geochem. Geophys. Geosyst. **16**, 1378–1400, 2015.
14. T. S. Baumann, B. J. P. Kaus, A. A. Popov, *Constraining effective rheology through parallel joint geodynamic inversion*, Tectonophysics **631**, 197–211, 2014.
15. T. S. Baumann, B. J. P. Kaus, *Geodynamic inversion to constrain the non-linear rheology of the lithosphere*, Geophys. J. Int. **202**, 1289–1316, 2015.
16. Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, Journal of large-scale research facilities, **1**, A1, 2015, `http://dx.doi.org/10.17815/jlsrf-1-18`.